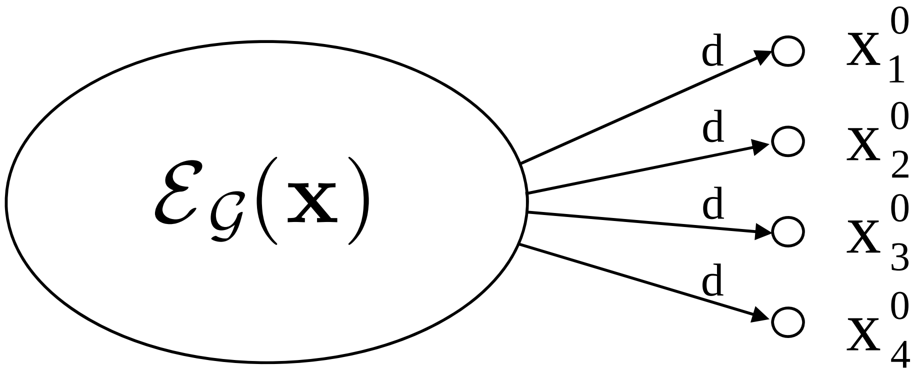
Klassifikation durch direkten Vergleich (Matching)

Eine triviale Lösung für die Klassifikation ergibt sich durch direkten Vergleich des unbekannten Musters in allen Erscheinungsformen der Äquivalenzklasse (Schablonenvergleich) gemäß einer Metrik d (Maß für die Klassenzugehörigkeit) mit den Prototypen einer Klasse (Bsp.: Holzpuzzle)

Äquivalenzklasse: $$\mathcal{E}_{\mathcal{G}}(\mathbf{x}) := \{ g_i(\mathbf{x}) \mid \forall g_i \in \mathcal{G} \}$$



Einführung einer Metrik d(**x**,**y**) für den Vergleich in unitären oder Innenprodukträumen (oder allgemeiner in Hilbert-Räumen): eine Metrik wird über das Innen- oder Skalarprodukt induziert

$$d(\mathcal{E}_{\mathcal{G}}(\mathbf{x}),\mathbf{x}_j^0) = \left\|\mathcal{E}_{\mathcal{G}}(\mathbf{x})-\mathbf{x}_j^0\right\| = <\mathcal{E}_{\mathcal{G}}(\mathbf{x})-\mathbf{x}_j^0,\mathcal{E}_{\mathcal{G}}(\mathbf{x})-\mathbf{x}_j^0>^{1/2}$$

$$= \sqrt{\left\|\mathcal{E}_{\mathcal{G}}(\mathbf{x})\right\|^2+\left\|\mathbf{x}_j^0\right\|^2-2<\mathcal{E}_{\mathcal{G}}(\mathbf{x}),\mathbf{x}_j^0>} = d_j(\mathbf{x}(\mathbf{p}),\mathbf{x}_j^0) \quad \text{für } \forall\mathbf{p}$$

Aus einem kleinsten Abstand resultiert ein größtes Innenprodukt bei normierten Vektoren (Maß für „Ähnlichkeit“)

Skalarprodukte für kontinuierliche und digitalisierte Muster:

Kontinuierliche Muster:

$$\text{1D:} \quad <\mathbf{x},\mathbf{y}>=\int x(t)y(t)dt$$

$$\text{2D:} \quad <\mathbf{X},\mathbf{Y}>=\iint X(t_1,t_2)Y(t_1,t_2)dt_1dt_2$$

Digitalisierte Muster:

$$\text{1D:} \quad <\mathbf{x},\mathbf{y}>=\sum x_i y_i$$

$$\text{2D:} \quad <\mathbf{X},\mathbf{Y}>=\sum\sum X_{i,j}Y_{i,j}$$

Berechnung des Skalarproduktes über den gesamten Parameterraum $\{\mathbf{p}\}$ der Bewegungsgruppe entspricht der Berechnung der Kreuzkorrelationsfunktion $R_{\mathrm{x,y}}$

$$< \mathcal{E}_{\mathcal{G}}(\mathbf{x}), \mathbf{x}_j^0 > = < \mathbf{x}(\mathbf{p}), \mathbf{x}_j^0 > = R_{\mathbf{x}(\mathbf{p}),\mathbf{x}_j^0} \quad \text{für } \forall \mathbf{p}$$

z.Bsp. Translation bei Bildern der Dimension $N \times N$ ergibt sich ein Aufwand, der proportional ist zu:

$$\underbrace{N^2}_{Skalarprodukt} \cdot \underbrace{N^2}_{zykl.Permut.} = N^4$$

Die Korrelation lässt sich mit der schnellen Fouriertransformation effizienter berechnen, nämlich mit einem Aufwand von $O(N^2\, ld\, N)$ gemäß (siehe DBV-I):

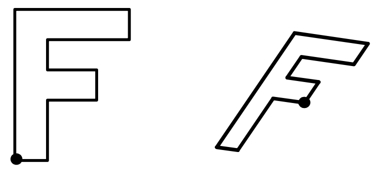
$$R_{x,y} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) \cdot \mathcal{F}^*(\mathbf{y}))$$

Berechnung des Skalarproduktes über den gesamten Parameterraum $\{\mathbf{p}\}$ der Bewegungsgruppe hat im allgemeinen für ein Signal oder Bild der Dimension N einen Aufwand von:

$$O(\underbrace{N}_{\text{Aufwand für ein Skalarprodukt}} \cdot \underbrace{N^{\dim(\mathbf{p})}}_{\text{Größe der Äquivalenzklasse}})$$

Hierbei geht man davon aus, dass man jede Achse des Parameterraumes mit der Bilddimension auflöst, also z.B. N Translationen.

Affin-invariante Erkennung mit Korrelation in allen Parametern (matched Filter)



Gegeben: Kontur mit N Punkten

➔ $O(N \times N^{\dim(\mathbf{p})}) = O(N^8)$ ops   mit $\dim(\mathbf{p})$: Freiheitsgrade

e.g. $N = 10^3 \Rightarrow \sim 10^{24}$ ops

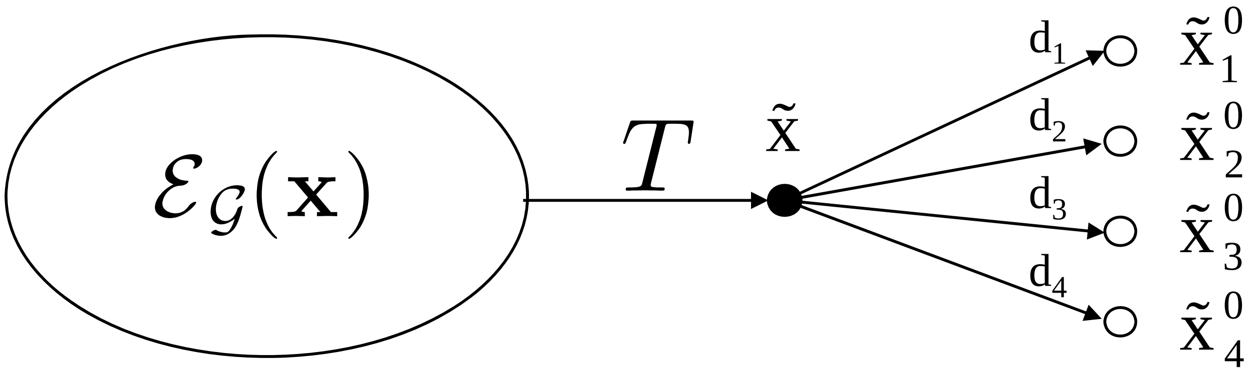
Annahme: 100 MIPS-Maschine (multiply-add)

$$\Rightarrow \sim \left(\frac{10^{24}\ \text{ops}}{10^8\ \text{ops/sec}}\right) = 10^{16}\ \text{sek} = 3.2 \times 10^8\ \text{Jahre}$$

Alter des Kosmos: 5×10^9 Jahre

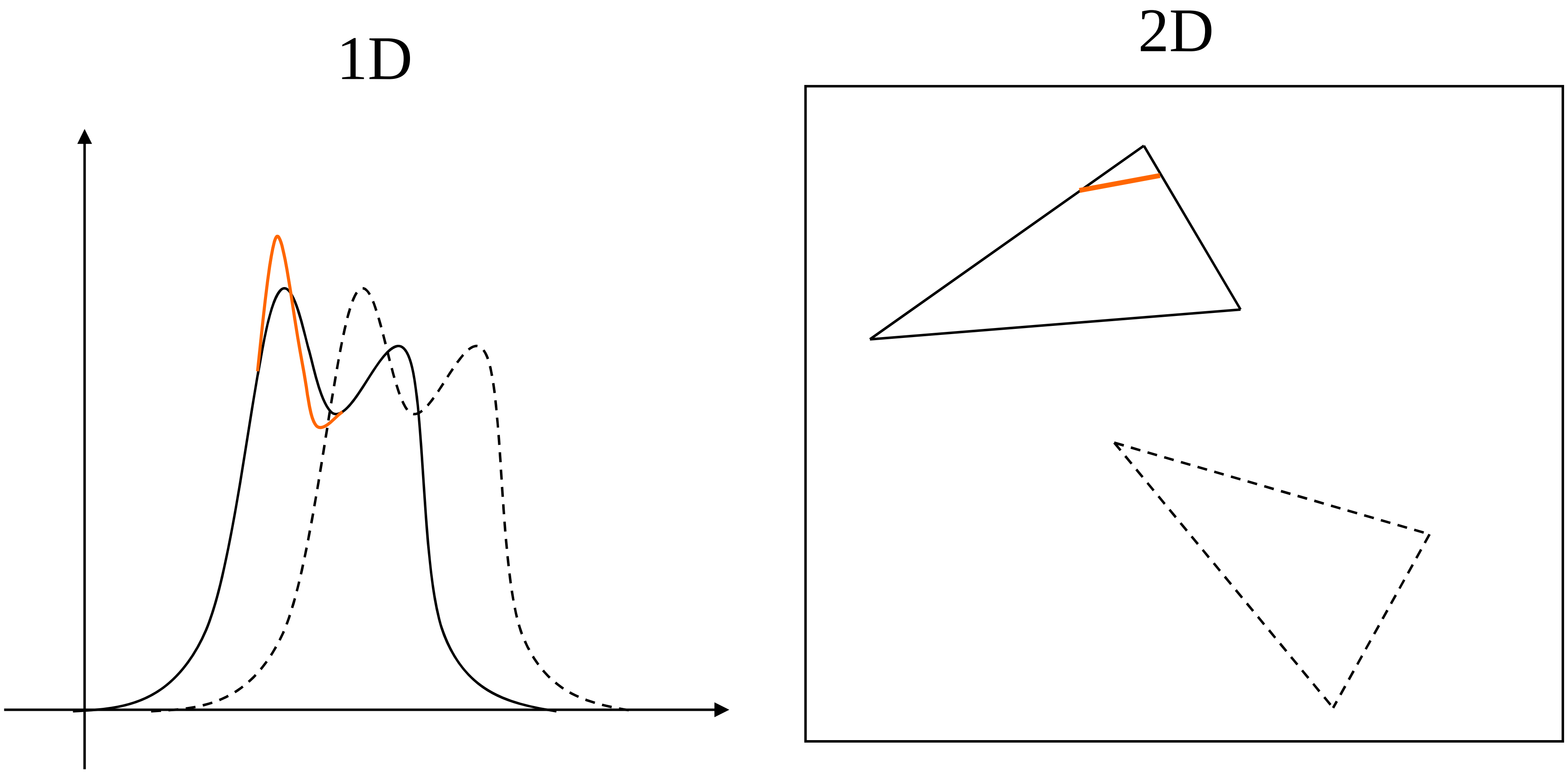
Klassifikation mit invarianten Merkmalen:

Extraktion lageinvarianter Merkmale, anschließend Klassifikation im Merkmalsraum:



$$\boxed{\tilde{\mathbf{x}} = T(\mathbf{x})}$$

Relevante and irrelevante Änderungen in Signalen und Bildern



Idee: entferne alle irrelevante Änderungen mit Hilfe einer mathematischen Transformation T und verwende nur invariante Merkmale

Notwendige Bedingung für die Invarianz einer Abbildung T :

1) Notwendige Bedingung für die Invarianz bezüglich der Gruppenwirkung $\mathcal{G}$**:**

$$\boxed{\mathbf{x}_1 \overset{\mathcal{G}}{\sim} \mathbf{x}_2 \;\Rightarrow\; T(\mathbf{x}_1) = \tilde{\mathbf{x}}_1 = \tilde{\mathbf{x}}_2 = T(\mathbf{x}_2)}$$

Alle Elemente einer Äquivalenzklasse werden in einen Punkt des Merkmalsraumes abgebildet. Dies kann auch wie folgt formuliert werden:

$$\boxed{T(g_i \mathbf{x}) = T(\mathbf{x}) \quad \text{für} \quad \forall g_i \in \mathcal{G}}$$

Invarianzbegriff bei zeitinvarianten Systemen

Der oben eingeführte Invarianzbegriff darf nicht verwechselt werden mit der Definiton *zeitinvarianter* Systeme. Dort verlangt man, dass die Systemantwort auf eine verschobene Erregung bei unveränderter ***Form*** ebenfalls verschoben ist. Dies gilt sowohl für die Erregung mit einem δ-Impuls, als auch für eine beliebige Erregung x. D.h.:

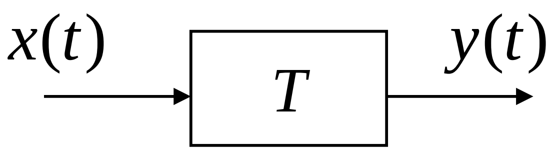
$$\delta(t) \overset{T}{\Rightarrow} h(t) \qquad \text{Impulsantwort}$$

$$\delta(t-t_0) \overset{T}{\Rightarrow} h(t-t_0) \qquad \text{verschobene Impulsantwort}$$

bzw.:

$$x(t) \overset{T}{\Rightarrow} y(t)$$

$$x(t-t_0) \overset{T}{\Rightarrow} y(t-t_0)$$



Operatorhomomorphismus

Bei zeitinvarianten Systemen handelt es sich um einen Operatorhomomorphismus, d.h.es gilt:

$$T(g_i(x(t))) = g_i(T(x(t))) = g_i(y(t))$$

wobei.: $$g_i(x(t)) = x(t - t_i)$$

Dies bedeutet, dass die Operationen g und T vertauscht werden dürfen!

Vollständigkeit einer invarianten Abbildung T :

2) Die ***Vollständigkeit*** ist eine hinreichende Bedingung dafür, dass aus der Gleichheit zweier Merkmalsvektoren darauf geschlossen werden darf, dass die beiden Objekte aus der gleichen Äquivalenzklasse stammen (Umkehrschluss der notwendigen Bedingung):

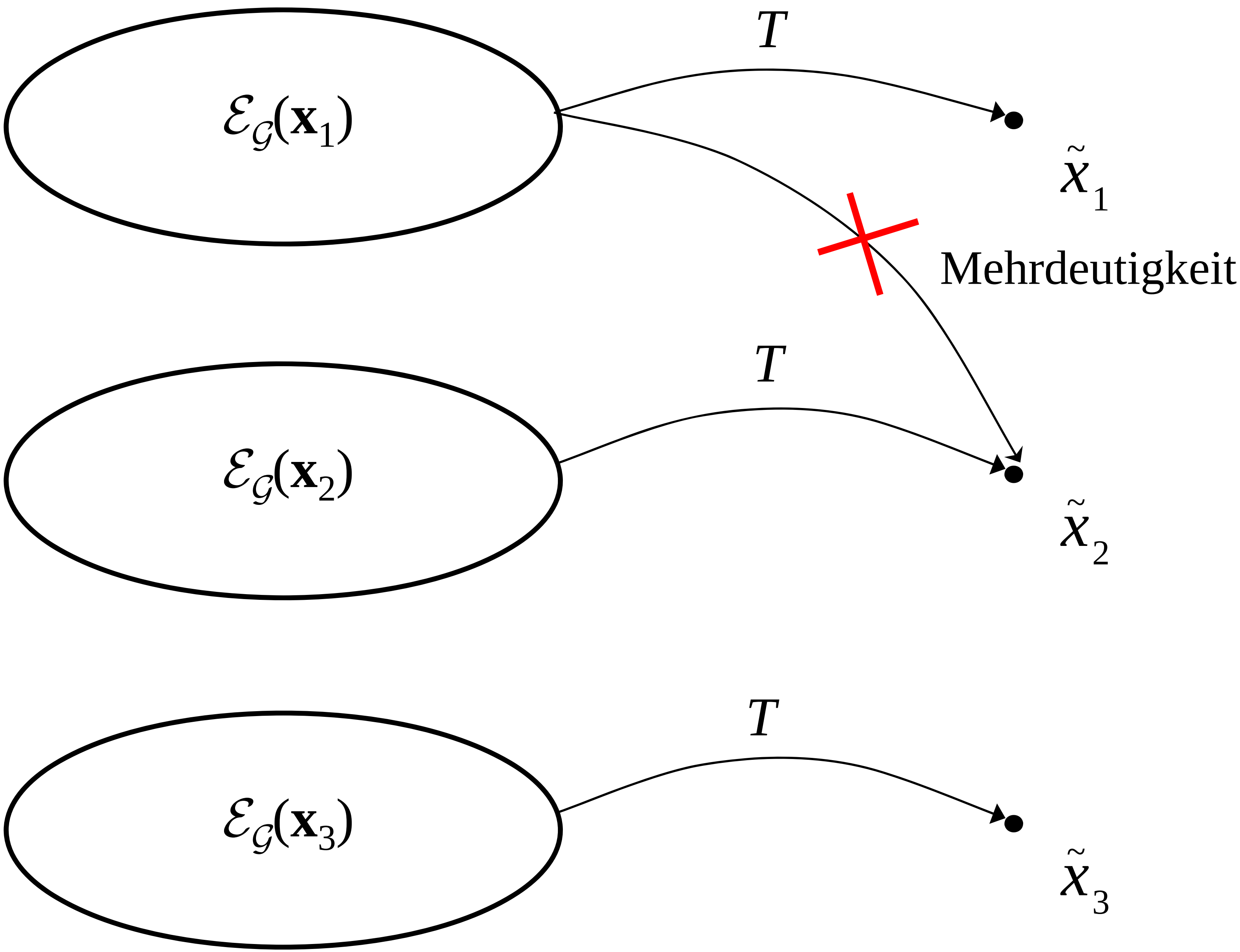
$$\boxed{T(\mathbf{x}_1) = T(\mathbf{x}_2) \Rightarrow \mathbf{x}_1 \overset{\mathcal{G}}{\sim} \mathbf{x}_2}$$

D.h. es gibt keine Mehrdeutigkeiten bei der Zuordnung von Merkmalsvektoren zu Äquivalenzklassen.

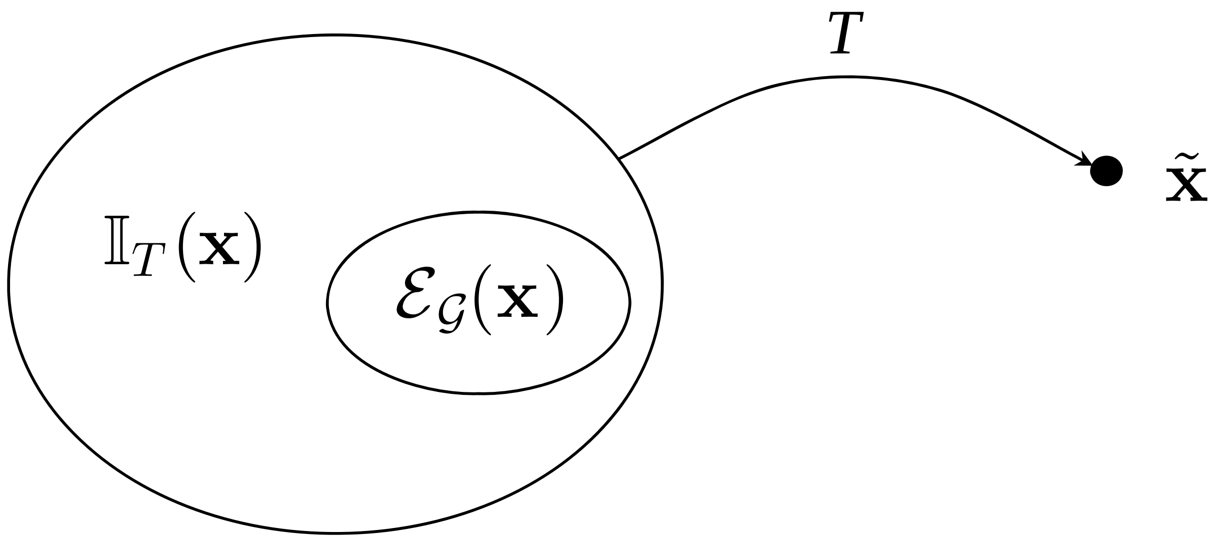
Schwächere Bedingung: ***Separierbarkeit*** *(Vollständigkeit für eine Teilmenge)*

also z.Bsp. Für die Menge aller Buchstaben bei der Zeichenerkennung

Abbildungseigenschaft der Transformation T:



Abbildungseigenschaften
Maß für den Grad der Vollständigkeit



1. Definition der Invarianten einer beliebigen kontrahienden Abbildung T:

$$\mathbb{I}_T(\mathbf{x}) := \{\mathbf{x_i} \mid T(\mathbf{x_i}) = T(\mathbf{x})\}$$

Abbildungseigenschaften

2. Aus der notwendigen Bedingung der Invarianz folgt:

$$\mathcal{E}_{\mathcal{G}}(\mathbf{x}) \subseteq \mathbb{I}_T(\mathbf{x}) \qquad \text{muss Teilmenge sein!}$$

3. Aus der Vollständigkeit folgt:

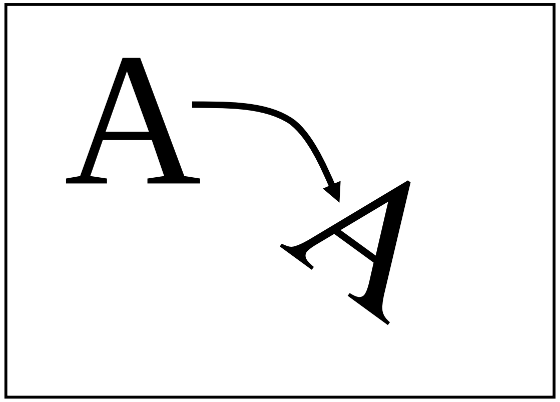
$$\mathcal{E}_{\mathcal{G}}(\mathbf{x}) = \mathbb{I}_T(\mathbf{x}) \qquad \forall \mathbf{x} \in \mathbb{X}$$

Bei der schwächeren Forderung der Separierbarkeit gilt diese Aussage nur für eine Teilmenge $\mathbb{X}_0$

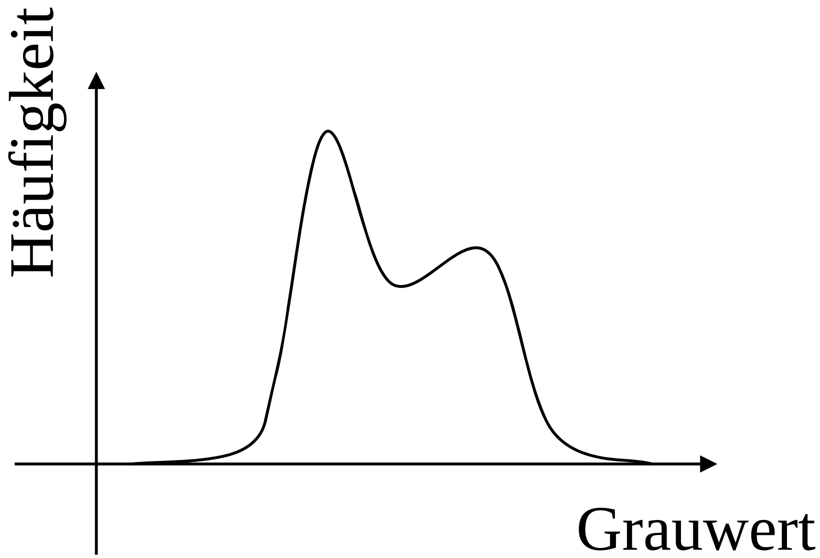
Beispiele für lageinvariante Merkmale

1. Der Mittelwert von Bildern ist invariant bezüglich Euklidscher Bewegungen (Translation und Rotation).
   Jedoch: Merkmale nicht vollständig!

$$\bar{x} = \frac{1}{N} \sum \sum x_{i,j}$$



2. Das Histogramm oder die Häufigkeitsverteilung ebenso. Auch hier ist die notwendige Bedingung erfüllt, die Merkmale sind jedoch stark mehrdeutig.

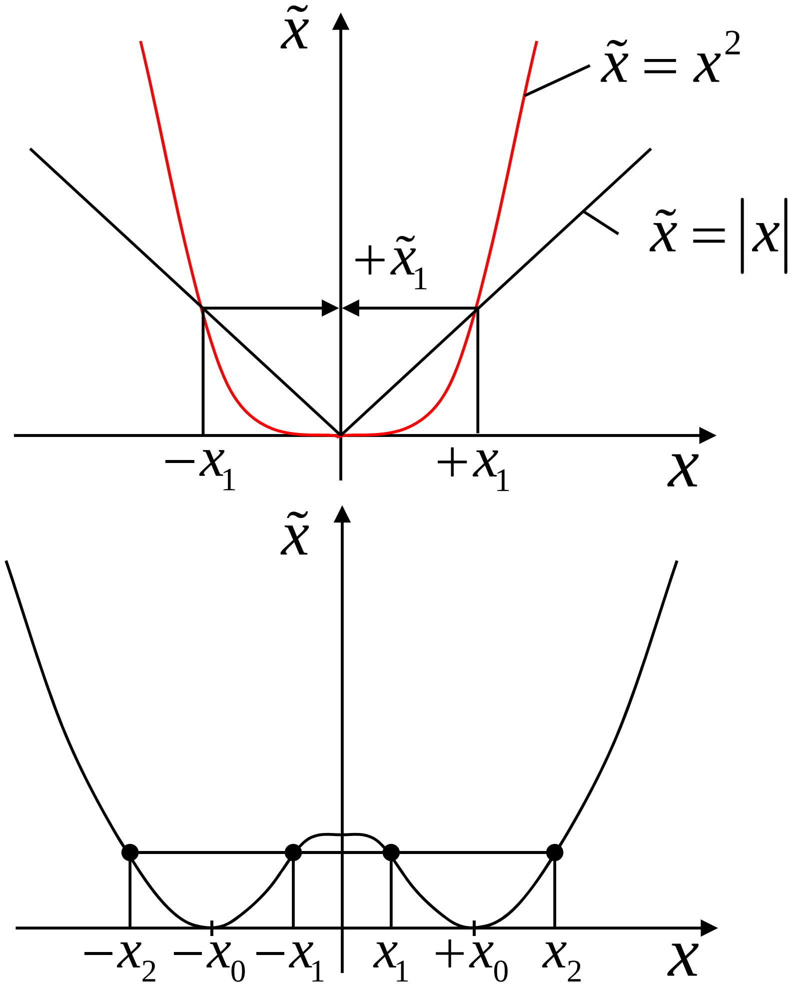


Man kann nämlich alle Bildpunkte beliebig permutieren, ohne das Histogramm zu verändern.

Das Histogramm ist vollständig für die symmetrische Gruppe (alle *N!* Permutationen).

Satz: Nur nichtlineare Abbildungen sind in der Lage von kompakten Äquivalenzklassen Invarianten zu bilden.

Einfache Beispiele für invariante Abbildungen mit nichtlinearen Funktionen



Die Äquivalenzklasse wird gebildet auf der Basis betrags-gleicher Zahlen, also z.Bsp.:

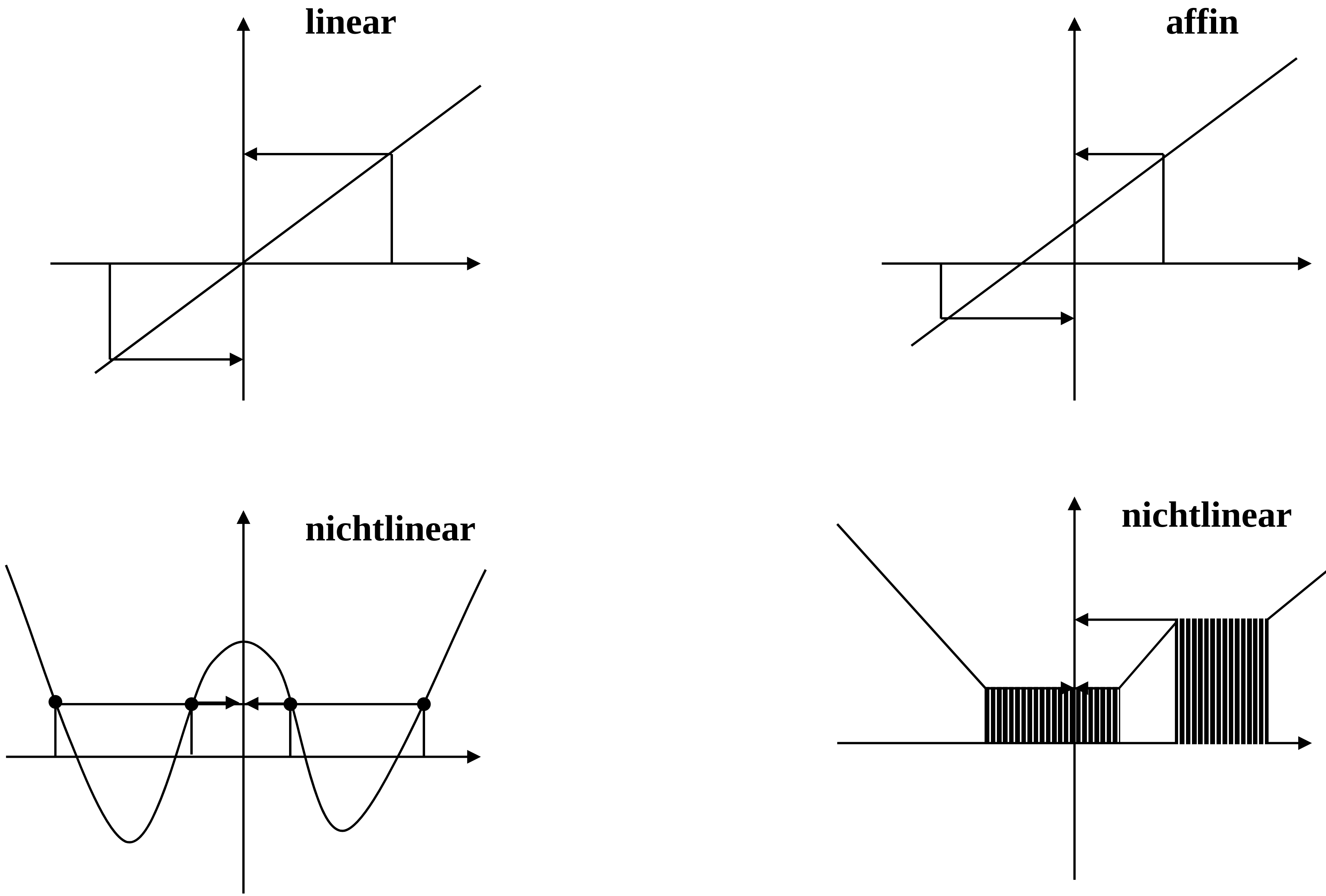
$$(-3) \overset{|\bullet|}{\sim} (+3)$$

Vollständige Invarianten!

$$\tilde{x} = a_0 + a_2 x^2 + a_4 x^4$$

Erfüllt ***notwendige***, aber nicht hinreichende Bedingung; Mehrdeutigkeiten. Aber z.Bsp. ***Separierbarkeit*** für $|x| < x_0$

Nichtlineare Abbildungen zur Erzeugung von Invarianten



Weitere Anforderungen an die invarianten Abbildungen

- Stetigkeit der Abbildung
- Erhaltung von Clustern
- Was passiert bei systematischen Störungen? (Helligkeit- und Kontrastschwankungen)

$$\boxed{\mathbf{x} = k\,\mathbf{x}^0 + a\,\mathbf{u}} \qquad \text{mit}: \quad \mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Lineare Systeme können diskriminieren zwischen ”Frequenzen“ oder spektralen Komponenten,

nichtlineare Systeme sind in der Lage zwischen “Äquivalenzklassen“ zu unterscheiden.

- **Lemma:**
  Gegeben sei eine *lineare* Abbildung *S*. Zwei Signale x_1,x_2 können nur zueinander äquivalent sein, wenn und nur wenn die Differenz x_1-x_2 ein Element des Nullraums *N(S)* ist (dieser ist jedoch nicht kompakt).
- ....
- Es ist unmöglich, mit linearen Abbildungen nichttriviale kompakte Äquivalenzklassen zu realisieren!